DENSITIES FOR ROUGH DIFFERENTIAL EQUATIONS UNDER HÖRMANDER'S CONDITION

THOMAS CASS AND PETER FRIZ

Abstract. We consider stochastic differential equations $dY = V(Y)\,dX$ driven by a multidimensional Gaussian process X in the rough path sense. Using Malliavin Calculus we show that Y_t admits a density for $t \in (0,T]$ provided (i) the vector fields $V = (V_1, ..., V_d)$ satisfy Hörmander's condition and (ii) the Gaussian driving signal X satisfies certain conditions. Examples of driving signals include fractional Brownian motion with Hurst parameter $H > 1/4$, the Brownian Bridge returning to zero after time T and the Ornstein-Uhlenbeck process.

1. Introduction

In the theory of stochastic processes, Hörmander's theorem on hypoellipticity of degenerate partial differential equations has always been an important tool to decide whether or not a diffusion process with given generator admits a density. This dependence on PDE theory was removed when P. Malliavin devised a purely probabilistic approach to Hörmander's theorem which is perfectly adapted to prove existence and smoothness of densities for diffusions given as strong solution to an Itô stochastic differential equation driven by Brownian motion.

The key ingredients of Malliavin's machinery, better known as *Malliavin Calculus* or *stochastic calculus of variations* can be formulated in the setting of an abstract Wiener space $(W, \mathcal{H}, \mu)$. This concept is standard (e.g. [30] or any modern book on stochastic analysis) as is the notion of *weakly non-degenerate* $\mathbb{R}^e$-valued functional φ which has the desirable property that the image measure $\varphi_* \mu$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^e$. (Functionals which are *non-degenerate* have a smooth density.) Precise definitions are given later on in the text.

Given these abstract tools, we turn to the standard Wiener space $C\left([0,T], \mathbb{R}^d\right)$ equipped with Wiener measure i.e. the standard model for Brownian motion $B = B(\omega)$. From Itô's theory, we know how to solve the stochastic differential equation

$$dY = \sum_{i=1}^{d} V_i(Y) \circ dB^i \equiv V(Y) \circ dB, \;\; Y(0) = y_0 \in \mathbb{R}^e.$$

The Itô-map $B \mapsto Y$ is notorious for its lack of strong regularity properties. On the positive side, it is smooth in a weak Sobolev type sense ("smooth in Malliavin's sense") and under Hörmander's condition at $y_0 \in \mathbb{R}^e$

$$(H): \quad \mathrm{Lie}\,[V_1, ..., V_d]_{y_0} = \mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e \tag{1.1}$$

one can show (e.g. [30, 36, 2, 32]) that the solution map $B \mapsto Y_t$ is non-degenerate for all $t \in (0,T]$. This line of reasoning provides a direct probabilistic approach to

the study of transition densities of Y and has found applications from stochastic fluid dynamics to interest rate theory, e.g. [12, 21]. The same range of applications[1] nowadays demand stochastic models of type

$$dY = V(Y)\, dX \tag{1.2}$$

where X is a Gaussian process, such as fractional Brownian motion (short: fBm). Differential equations of this type have also been used as simple examples for the study of ergodicity of non-Markovian systemems, [24].

In a previous paper [4] we linked rough paths and Malliavin calculus by giving a (simple) proof that existence of a density for solutions of (1.2) holds true under ellipticity i.e.

$$(E): \quad \mathrm{Span}\,[V_1, ..., V_d]_{y_0} = \mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e.$$

and generic non-degeneracy conditions on X, the differential equation (1.2) being understood in the rough path sense [25, 29], a unified framework which covers *at once* Young and Stratonovich solutions (and goes well beyond). The aim of this paper is to prove the existence of densities under Hörmander's condition (H) in the following form:

Theorem 1. *Let $\left(X_t^1, \dots, X_t^d\right) = (X_t : t \in [0,T])$ be a continuous, centered Gaussian process with independent components $X^1, \dots, X^d$. Assume X satisfies the conditions listed in section 4. (In particular, X is assumed to lift to a geometric rough path so that (1.2) makes sense as random rough differential equations.) Let $V = (V_1, ..., V_d)$ a collection of smooth bounded vector fields on $\mathbb{R}^e$ with bounded derivatives which satisfies Hörmander's condition (H) at y_0. Then the random RDE solution $Y_t = Y_t(\omega) \in \mathbb{R}^e$ to (1.2) started at $Y_0 = y_0$ admits a density with respect to Lebesgue measure on $\mathbb{R}^e$ for all times $t \in (0,T]$.*

One should note that X, the Gaussian driving signal of (1.2), is fully described by the covariance function of each component and, under the further assumption of IID components, by the covariance of a single component, i.e. $R(s,t) = \mathbb{E}\left(X_s^1 X_t^1\right)$. In principle all conditions on X are checkable from the covariance, in practice it is convenient to have conditions available which involve the *reproducing kernel Hilbert* $\sim$ or *Cameron-Martin space* associated to X as well as certain sample path properties. Leaving these technical details to section 4 we emphasize that our conditions are readily checked in many cases and now give a list of examples to which our theorem applies. It may be helpful to note that whenever X is a semi-martingale on $[0,T]$ then (1.2) can be understood as Stratonovich stochastic differential equation, i.e.

$$dY = \sum_{i=1}^{d} V_i(Y) \circ dX^i.$$

In such cases, rough path theory appears as intermediate tool that is neither needed to understand the assumptions nor the conclusions of theorem 1. There may be cases when X can be written in terms of Brownian motion so that ultimately the techniques of [5, 38] are applicable. But in general theorem 1 covers new grounds.

---

[1] For instance, stochastic differential equations driven by fBM have applications applications to vortex filaments; applications to finance include (geometric) fractional Brownian motion as paradigm of a non-semimartingale which admits no arbitrarge under transaction costs. The reader is refered to the books [32, Sec. 5.3 and 5.4], [9, Sec 8.1.] and the references therein.

Example 1 (Brownian motion)**.** *When $R(s,t) = \min(s,t)$ the driving signal $X = X(\omega)$ is a d-dimensional standard Brownian motion on $[0,T]$. As is well understood* [29] *one needs to add Lévy's area process to obtain a geometric rough path (known in this context as Brownian rough path or Enhanced Brownian motion). A solution to (1.2) in the rough path sense then precisely solves the stochastic differential equation in Stratonovich form*

$$dY = \sum_{i=1}^{d} V_i(Y) \circ dB^i.$$

Subject to Hörmander's condition (H), Theorem 1 then shows that $Y_t = Y_t(\omega)$ has a density for $t > 0$ which is of course well-known.

Example 2 (Fractional Brownian motion)**.** *When $2R(s,t) = t^{2H} + s^{2H} - |t-s|^{2H}$ the corresponding process is known as fractional Brownian motion $B^H = B^H(\omega)$ with Hurst parameter $H \in (0,1)$. Popularized by* [31] *it generalizes Brownian motion which corresponds to $H = 1/2$. For fixed H and when no confusion is possible we shall write $B = (B^1, \ldots, B^d)$ for d-dimensional fBm. When $H > 1/2$, Kolmogorov's criterion shows that B has nice sample paths (more precisely, Hölder continuous sample paths of exponent greater $1/2$) which has the great advantage that (1.2) can be understood as integral equation for fixed ω based on Young integrals (i.e. limits of Riemann-Stieltjes sums). In this setting of nice sample paths, existence of a density was established in* [33] *assuming ellipticity. Using deterministic estimates for the Jacobian of the flow this density was then shown to be smooth* [34]*; building on the same estimates the Hörmander case was obtained in* [1]*. For $H \leq 1/2$ the situation appears to be fundamentally different: first, in view of Brownian (and worse) sample path regularity one needs Itô or rough path ideas to make sense of (1.2) for $H \leq 1/2$. Secondly, the proof of* [1] *does not extend to the rough path setting*[2] *and relies somewhat delicately on specific properties of fractional Brownian motion. In any case, theorem 1 shows that $Y_t = Y_t(\omega)$, solution to the RDE driven by multidimensional fBM with Hurst parameter $H > 1/4$ has a density for all positive times provided the vector fields satisfy Hörmander's condition*[3]*. The novelity is of course the degenerate regime $H < 1/2$ with sample path regularity worse than Brownian motion.*

Example 3 (Brownian Bridge)**.** *Let B be a d-dimensional standard Brownian motion. Define the Brownian bridge returning to zero at time T by*

$$X_t^T := B_t - \frac{t}{T} B_T \text{ for } t \in [0,T].$$

Equivalently, one can define X^T via the covariance

$$R^T(s,t) = \min(s,t)(1 - \max(s,t)/T).$$

[2] The estimates of [34] can be generalized [20] to (sharp) deterministic estimates on the Jacobian of RDE solutions giving L^p-estimates on the flow of RDEs driven by fBM if and only if $H > 1/2$. In particular, one sees that L^p-estimates on the flow of Stratonovich SDEs ($H = 1/2$) are fundamentally probabilistic i.e. rely on cancellations in stochastic integration. At present, the question of how to obtain good integrability when $H < 1/2$ is open although one suspects that Gaussian isoperimetry will ultimately play a role.

[3] As is well understood [29], for $H \leq 1/4$ fractional Brownian increments decorrelate too slowly for stochastic area to exist and so there is no meaningful lift of fBM with $H \leq 1/4$ to a geometric rough path.

Clearly $X_t^T|_{t=T} = 0$ and trivially (take $dY = dX$) the conclusion of Theorem 1 cannot hold; this behaviour is indeed ruled out by condition 3 in section 4. On the other hand, we may consider $X^{T+\varepsilon}$ restricted to $[0,T]$ and in this case the conditions in section 4 are readily verified. It is worth remarking that $Z := X^{T+\varepsilon}$ stopped at time T is also a semi-martingale; for instance, by writing $\left(X_t^{T+\varepsilon} : t \leq T\right)$ as strong solution to an Itô differential equation with (well-behaved) drift (as long as $t \leq T$). The conclusion of theorem 1 can then be stated by saying that the unique Stratonovich solution to $dY = \sum V_i(Y) \circ dZ^i$ admits a density for all times $t \in (0,T]$ provided the vector fields satisfy Hörmander's condition (H).

Example 4 (Ornstein-Uhlenbeck)**.** *Let B be a standard d-dimensional Brownian motion and define the centered Gaussian process X by Wiener-Itô integration,*

$$X_t^i = \int_0^t e^{-(t-r)} dB_r^i \text{ with } i = 1, ..., d.$$

X satisfies the Itô differential equations, $dX_t = -X_t dt + dB_t$ and is also a semi-martingale. The conditions of section 4 are readily checked (in essence one uses $X_t \sim B_t$ at $t \to 0+$ and the absence of Brownian Bridge type degeneracy). The conclusion of theorem 1 can then be stated by saying that the unique Stratonovich solution to $dY = \sum V_i(Y) \circ dX^i$ admits a density for all positive times provided the vector fields satisfy Hörmander's condition (H).

Further examples (for instance, "fractional" version of the Brownian Bridge and Ornstein-Uhlenbeck process) are readily constructed. Generalizing examples 2, 4 one could consider Volterra processes [8], i.e. Gaussian process with representation $X = \int K(\cdot, s)\, dB_s$, and derive sufficient conditions on the kernel K which imply those of section 4. Existence of a rough path lift of X aside, one would need non-degeneracy of K and certain scaling properites as $t \to 0+$ but we shall not pursue this here. (In any case, there are non-Volterra examples, such as the Brownian bridge returning to zero at $(T + \varepsilon)$, to which theorem 1 applies.)

The proof of theorem 1 is based on the fact [4] that RDE solutions driven by Gaussian signals are "$\mathcal{H}$-differentiable" i.e. differentiable in Cameron-Martin directions. Existence of a density is then reduced to showing that the Malliavin covariance matrix is weakly non-degenerate. The standard proof of this (e.g. [30, 2] or [32, Sec 2.3.2]) is based on Blumenthal's 0-1 law and the Doob-Meyer decomposition for semi-martingales. The main difficulty to overcome in the general Gaussian context of this paper is that the Doob-Meyer decomposition is not available and we manage to bypass its use by suitable *small time developments for RDEs*, obtained in [19]; in conjunction with *(Stroock-Varadhan type) support description for certain Gaussian rough paths* (as conjectued by Ledoux et al. [28] and carried out independently in [11, 15], see also [4] and [18].)

The crucial induction step - which explains the appearance of higher brackets - requires us to assume a "non-standard" Hörmander condition which involves only iterated Lie-brackets contracted against certain tensors arising from free nilpotent Lie groups. Equivalence to the usual Hörmander condition (H) is then established separately.

2. Preliminaries on ODE and RDEs

2.1. **Controlled ordinary differential equations.** Consider the controlled ordinary differential equations, driven by a smooth R^d-valued signal $f = f(t)$ along sufficiently smooth and bounded vector fields $V = (V_1, ..., V_d)$,

$$dy = V(y)df \equiv \sum_{i=1}^{d} V_i(y) f'(t) \, dt, \; y(t_0) = y_0 \in \mathbb{R}^e. \tag{2.1}$$

We call $U^f_{t \leftarrow t_0}(y_0) \equiv y_t$ the associated flow. Let J denote the Jacobian of U. It satisfies the ODE obtain by formal differentiation w.r.t. y_0. More specifically,

$$a \mapsto \left\{ \frac{d}{d\varepsilon} U^f_{t \leftarrow t_0}(y_0 + \varepsilon a) \right\}_{\varepsilon=0}$$

is a linear map from $\mathbb{R}^e \to \mathbb{R}^e$ and we let $J^f_{t \leftarrow t_0}(y_0)$ denote the corresponding $e \times e$ matrix. It is immediate to see that

$$\frac{d}{dt} J^f_{t \leftarrow t_0}(y_0) = \left[\frac{d}{dt} M^f \left(U^f_{t \leftarrow t_0}(y_0), t \right) \right] \cdot J^f_{t \leftarrow t_0}(y_0)$$

where $\cdot$ denotes matrix multiplication and

$$\frac{d}{dt} M^f(y, t) = \sum_{i=1}^{d} V_i'(y) \frac{d}{dt} f^i_t.$$

Note that $J^f_{t_2 \leftarrow t_0} = J^f_{t_2 \leftarrow t_1} \cdot J^f_{t_1 \leftarrow t_0}$. We can also consider Gateaux derivates in the driving signal and define

$$D_h U^f_{t \leftarrow 0} = \left\{ \frac{d}{d\varepsilon} U^{f+\varepsilon h}_{t \leftarrow 0} \right\}_{\varepsilon=0}.$$

One sees that $D_h U^f_{t \leftarrow 0}$ satisfies a linear ODE and the variation of constants formula leads to

$$D_h U^f_{t \leftarrow 0}(y_0) = \int_0^t \sum_{i=1}^{d} J^f_{t \leftarrow s} \left(V_i \left(U^f_{s \leftarrow 0} \right) \right) dh^i_s.$$

Finally, given a smooth vector field W a straight-forward computation gives

$$dJ^f_{0 \leftarrow t} \left(W \left(U^f_{t \leftarrow 0} \right) \right) = J^f_{0 \leftarrow t} \left([V_i, W] \left(U^f_{t \leftarrow 0} \right) \right) df^i_t. \tag{2.2}$$

2.2. **Rough differential equations.** Following [25, 29, 17] a geometric p-rough path $\mathbf{x}$ over $\mathbb{R}^d$ is a continuous path on $[0, T]$ with values in $G^{[p]}(\mathbb{R}^d)$, the step-$[p]$ nilpotent group over $\mathbb{R}^d$, and of finite p-variation relative to the [17] Carnot-Caratheodory metric d on $G^{[p]}(\mathbb{R}^d)$, i.e.

$$\sup_{n \in \mathbb{N}} \sup_{0 < t_1 < \ldots < t_n < T} \sum_i d\left(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}} \right)^p < \infty.$$

As in [6, 25] we view $G^{[p]}(\mathbb{R}^d)$ as embedded in its enveloping tensor algebra i.e.

$$G^{[p]}(\mathbb{R}^d) \subset T^{[p]}(\mathbb{R}^d) := \oplus_{k=0,\ldots,[p]} \left(\mathbb{R}^d \right)^{\otimes k}.$$

One can then think of $\mathbf{x}$ as a path $x : [0,T] \to \mathbb{R}^d$ enhanced with its iterated integrals although the later need not make classical sense[4]. The canonical projection to $(\mathbb{R}^d)^{\otimes k}$ is denoted $\pi_k(\mathbf{x})$ or $\mathbf{x}^k$. Lyons' theory of rough paths then gives deterministic meaning to the rough differential equation (short: RDE)

$$dy = V(y)\, d\mathbf{x}. \tag{2.3}$$

(One can think of RDE solutions as limit points of corresponding ODEs of form (2.1) in which the smooth driving signals *plus their iterated integrals up to order* $[p]$ converge to $\mathbf{x}$ in suitable p-variation distance.) The motivating example, e.g. [25, 29], is that *almost every* continuous joint realization of Brownian motion and Lévy's area process (equivalently: iterated Stratonovich integrals) gives rise to a geometric p-rough path for $p > 2$, known as Brownian rough path or Enhanced Brownian motion (cf. example 1) which provides in particular a robust path-by-path view of Stratonovich SDEs.

Back to the deterministic RDE (2.3) and assuming smoothness of the vector fields $V = (V_1, ..., V_d)$, the solution induces a flow $y_0 \mapsto U^{\mathbf{x}}_{t \leftarrow t_0}(y_0)$. Following [26, 27], the Jacobian $J^{\mathbf{x}}_{t \leftarrow t_0}$ of the flow exists and satisfies a linear RDE, as does the directional derivative

$$D_h U^{\mathbf{X}}_{t \leftarrow 0} = \left\{ \frac{d}{d\varepsilon} U^{T_{\varepsilon h}\mathbf{x}}_{t \leftarrow 0} \right\}_{\varepsilon = 0}$$

for a smooth path h. If $\mathbf{x}$ arises from a smooth path x together with its iterated integrals the *translated rough path* $T_h \mathbf{x}$ (cf. [26, 29]) is nothing but $x + h$ together with its iterated integrals. In the general case, we assume $h \in C^{q\text{-var}}$ with $1/p + 1/q > 1$, the translation $T_h \mathbf{x}$ can be written in terms of $\mathbf{x}$ and cross-integrals between $\pi_1(\mathbf{x}_{0,\cdot}) =: x$ and the perturbation h. (These integrals are well-defined Young-integrals.)

Proposition 1. *Let* $\mathbf{X}$ *be a geometric* p*-rough paths over* $\mathbb{R}^d$ *and* $h \in C^{q\text{-}var}([0,T], \mathbb{R}^d)$ *such that* $1/p + 1/q > 1$. *Then*

$$D_h U^{\mathbf{X}}_{t \leftarrow 0}(y_0) = \int_0^t \sum_i J^{\mathbf{X}}_{t \leftarrow s}\left(V_i\left(U^{\mathbf{X}}_{s \leftarrow 0}\right)\right) dh^i_s$$

where the right hand side is well-defined as Young intergral.

Proof. $J^{\mathbf{X}}_{t \leftarrow 0}, D_h U^{\mathbf{X}}_{t \leftarrow 0}$ satisfy (at least jointly with $U^{\mathbf{X}}_{t \leftarrow 0}$) RDEs driven by $\mathbf{X}$ which allows, in essence, to use Lyons' limit theorem; this is discussed in detail in [26, 27]. A little care is needed since the resulting vector fields are not bounded anymore. However we can rule out explosion and then localize the problem: the needed remark is that $J^{\mathbf{X}}_{t \leftarrow 0}$ also satisfy a linear RDE of form

$$dJ^{\mathbf{X}}_{t \leftarrow 0} = dM^{\mathbf{X}}\left(U^{\mathbf{X}}_{t \leftarrow 0}(y_0), t\right) \cdot J^{\mathbf{X}}_{t \leftarrow 0}(y_0)$$

and explosion can be ruled out by direct iterative expansion and estimates of the Einstein sum as in [25]. □

[4] In fact, $G^N(\mathbb{R}^d)$ can realized as all points in the tensor algebra which arise from computing iterated integrals up to order N of smooth paths over a fixed time interval. The group product then corresponds to the concatenation of paths, the inverse corresponds to running a path backwards in time etc.

3. RDEs driven by Gaussian signals

We consider a continuous, centered Gaussian process $X = \left(X^1, ..., X^d\right)$ with independent components started at zero. This gives rise to an abstract Wiener space $(W, \mathcal{H}, \mu)$ where $W = \bar{\mathcal{H}} \subset C_0\left([0,T], \mathbb{R}^d\right)$. Note that $\mathcal{H} = \oplus_{i=1}^d \mathcal{H}^{(i)}$ and recall that element of $\mathcal{H}$ are of form $h_t = \mathbb{E}\left(X_t \xi(h)\right)$ where $\xi(h)$ is a Gaussian random variable. The ("reproducing kernel") Hilbert-structure on $\mathcal{H}$ is given by $\langle h, h' \rangle_{\mathcal{H}} := \mathbb{E}\left(\xi(h)\xi(h')\right)$.

Existence of a Gaussian geometric p-rough path above X is tantamount to the existence of certain Lévy area integrals. The case of fractional Brownian motion is well understood and several construction have been carried out [7, 29, 11, 35]. In particular, one requires $H > 1/4$ for the existence of stochastic areas (which can be defined as $L^2(\mathbb{P})$-limits as in Itô's theory). Resultingly, one has to deal with geometric p-rough paths for $p < 4$. (When $p < 2$ there is enough sample path regularity to use Young integration and we avoid speaking of rough paths.)

Condition 1. *Assume X lifts to a (random) geometric p-rough path $\mathbf{X}$ and $\exists q : 1/p + 1/q > 1$ such that*

$$\mathcal{H} \hookrightarrow C^{q\text{-}var}\left([0,T], \mathbb{R}^d\right).$$

The example to have in mind is Brownian motion for which the above condition is satisfied with $p = 2 + \varepsilon$ and $q = 1$. (We shall say more about other Gaussian examples in section 4.)

If $\mathbf{X} = \mathbf{B}^H$ denotes the geometric p-rough path, $p \in (1/H, [1/H] + 1)$, associated to fractional Brownian motion then it satisfies a *Stroock-Varadhan support description in rough path topology.* This was first conjectued by Ledoux et al. [28] (who obtained it for the Brownian rough path) and carried out independently in [11, 15] for $H > 1/3$. The case of $H > 1/4$ is more difficult and discussed in some detail in [4]. A proof in the generic context of Gaussian rough paths (covering fBM with $H > 1/4$ as special case) is given in the forthcoming paper [18]. The statement is

$$\operatorname{supp}\left(\mathbb{P}_* \mathbf{X}\right) = \overline{\left\{S_{[p]}\left(\mathcal{H}\right)\right\}} \tag{3.1}$$

where support and closure are relative to the homogenous p-variation topology for geometric p-rough paths. We recall that $S_{[p]}$, for $[p] = 2, 3$ given by

$$\begin{aligned} S_2 &: \quad h \mapsto 1 + \int_0^t dh + \int_0^t \int_0^s dh \otimes dh \\ S_3 &: \quad h \mapsto 1 + \int_0^t dh + \int_0^t \int_0^s dh \otimes dh + \int_0^t \int_0^s \int_0^r dh \otimes dh \otimes dh \end{aligned}$$

lifts $\mathbb{R}^d$-valued paths canonically to $G^{[p]}\left(\mathbb{R}^d\right)$-valued paths. In [18] it is seen that $\mathbf{X}$ exists provided the covariance has finite ρ-variation with $\rho < 2$ and it is also established that $\mathcal{H} \hookrightarrow C^{\rho\text{-var}}$ which guarantees that $S_{[p]}\left(\mathcal{H}\right)$ is well-defined via Young integration. Such support description will be important in checking condition 5, section 4.

Definition 1. [36, 32, 30] *Given an abstract Wiener space $(W, \mathcal{H}, \mu)$, a r.v. $F : W \to \mathbb{R}$ is $\mathcal{H}$ differentiable at $\omega \in W$ iff exists $DF(\omega) \in \mathcal{H}^*$ such that*

$$\forall h \in \mathcal{H} : \left\{ \frac{d}{d\varepsilon} F\left(\omega + \varepsilon h\right) \right\}_{\varepsilon = 0} = \langle DF(\omega), h \rangle_{\mathcal{H}}.$$

A vector-valued r.v. $F = \left(F^1, ..., F^e\right) : W \to \mathbb{R}^e$ *is* $\mathcal{H}$ *differentiable iff each* F^i *is* $\mathcal{H}$ *differentiable. In this case,* $DF(\omega) = \left(DF^1(\omega), ..., DF^e(\omega)\right)$ *is a linear bounded map from* $\mathcal{H} \to \mathbb{R}^e$*. One then defines the Malliavin covariance matrix as the random matrix*

$$\sigma(\omega) := \left(\left\langle DF^i, DF^j \right\rangle_{\mathcal{H}}\right)_{i,j=1,\dots,e} \in \mathbb{R}^{e\times e}.$$

We call F *weakly non-degenerate if* $\det(\sigma) \neq 0$ *almost surely.*

Proposition 2. *Assume condition (1). Then, for fixed* $t \geq 0$*, the* $\mathbb{R}^e$*-valued random variable*

$$\omega \mapsto U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)$$

is almost surely $\mathcal{H}$*-differentiable.*

Proof. By assumption $1/p + 1/q > 1$. We may assume that $\mathbf{X}(\omega)$ has been defined so that $\mathbf{X}(\omega)$ is a geometric p-rough path for every $\omega \in W$. We also know that $\mathbf{X}(\omega + h) \equiv T_h\mathbf{X}$ almost surely and so

$$\mathbb{P}\left[\mathbf{X}(\omega + \varepsilon_n h) \equiv T_{\varepsilon h}\mathbf{X} \text{ for any countable family } (\varepsilon_n)\right] = 1$$

We fix ω in the above set of full measure. For fixed t, define

$$Z_{i,s} \equiv J_{t\leftarrow s}^{\mathbf{X}}\left(V_i\left(U_{s\leftarrow 0}^{\mathbf{X}}\right)\right).$$

Noting that $s \mapsto Z_{i,s}$ is in $C^{p\text{-var}}$ we have, with implicit summation over i,

$$\begin{aligned}
\left|D_h U_{t\leftarrow 0}^{\mathbf{X}}(y_0)\right| &= \left|\int_0^\tau J_{t\leftarrow s}^{\mathbf{X}}\left(V_i\left(U_{s\leftarrow 0}^{\mathbf{X}}\right)\right) dh_\lambda^i\right| \\
&= \left|\int_0^\tau Z_i dh^i\right| \\
&\leq c\left(|Z|_{p-var} + |Z(0)|\right) \times |h|_{\rho-var} \\
&\leq c\left(|Z|_{p-var} + |Z(0)|\right) \times |h|_{\mathcal{H}}
\end{aligned}$$

(We used Young's inequality.) The linear map $DU_{t\leftarrow 0}^{\mathbf{X}}(y_0) : h \mapsto D_h U_{t\leftarrow 0}^{\mathbf{X}}(y_0) \in \mathbb{R}^e$ is bounded and each component is an element of $\mathcal{H}^*$, hence

$$h \mapsto \left\{\frac{d}{d\varepsilon} U_{t\leftarrow 0}^{T_{\varepsilon h}\mathbf{X}(\omega)}(y_0)\right\}_{\varepsilon=0} = \left\langle DU_{t\leftarrow 0}^{\mathbf{X}}(y_0), h\right\rangle_{\mathcal{H}}.$$

Noting that the derivative at $\varepsilon = 0$ exists, by definition, if the difference quotients converge as $\varepsilon \downarrow 0$ and this holds iff convergence to the same limit takes place along any sequence $\varepsilon_n \downarrow 0$. It follows that, for almost every ω,

$$h \mapsto \left\{\frac{d}{d\varepsilon} U_{t\leftarrow 0}^{\mathbf{X}(\omega+\varepsilon h)}(y_0)\right\}_{\varepsilon=0} = \left\langle DU_{t\leftarrow 0}^{\mathbf{X}}(y_0), h\right\rangle_{\mathcal{H}}$$

and so the random variable $U_{t\leftarrow 0}^{\mathbf{X}}(y_0)$ is indeed a.s. $\mathcal{H}$-differentiable. □

4. Conditions on Driving Process

We now give a complete list of assumptions on the (d-dimensional) Gaussian driving signal $(X_t : t \in [0, T])$. The first condition was already needed in the previous section to show $\mathcal{H}$-differentiability of RDE solutions driven by X; we repeat it for completeness and to give some additional examples.

Condition 2. *Assume X lifts to a (random) geometric p-rough path $\mathbf{X}$ and $\exists q$: $1/p+1/q>1$ such that*

$$\mathcal{H} \hookrightarrow C^{q\text{-}var}\left([0,T],\mathbb{R}^d\right). \tag{4.1}$$

In the Brownian motion case this holds, as already remarked earlier, with $p=2+\varepsilon$ and $q=1$. The same is true for the Brownian bridge and the Ornstein-Uhlenbeck examples discussed in the introduction; although case-by-case verifications are not difficult, there is general criterion on the covariance which implies (4.1), see [17, Prop 16 applied with $\rho=1$], which also covers fBM. Let us give a direct argument for case of fBM which covers any Hurst parameter $H>1/4$. Writing $\mathcal{H}^H$ for the Cameron-Martin space of fBM, the variation embedding in [16] gives

$$\mathcal{H}^H \hookrightarrow C^{q\text{-var}} \text{ for any } q > (H+1/2)^{-1}.$$

At the same time [7, 29, 11, 35] fBM lifts to a geometric p-rough path for $p>1/H$. By choosing p,q small enough $1/p+1/q$ can be made arbitrarily close to $H+(H+1/2)=2H+1/2>1$ and so (4.1) holds indeed for fBM with Hurst parameter $H>1/4$.

Condition 3. *Fix $T>0$. We assume non-degeneracy on $[0,T]$ in the sense that for any smooth $f=(f_1,...,f_d):[0,T]\to\mathbb{R}^d$ we have*

$$\left(\int_0^T f dh \equiv \sum_{j=1}^d \int_0^T f_j dh^j = 0 \forall h \in \mathcal{H}\right) \implies f \equiv 0.$$

Again, fBM satisfies the following non-degeneracy condition simply because $C_0^1\left([0,T],\mathbb{R}^d\right) \subset \mathcal{H}^H$, cf. [15]. A Brownian bridge which returns to zero at time T is ruled out, while a Brownian bridge which returns to zero after time T is allowed. Checking condition 3 for the Ornstein-Uhlenbeck example in the introduction is left as exercise for the reader. The following lemma taken from [4] contains a few ramifications concerning condition 3; since $\mathcal{H}=\oplus_{i=1}^d \mathcal{H}^{(i)}$ there is no loss in generality in assuming $d=1$.

Lemma 1. *Assume*

$$1/p+1/q>1.$$

(i) The requirement that f is smooth above can be relaxed to $f\in C^{p\text{-}var}$.
(ii) The requirement that $\int f dh = 0 \forall h \in \mathcal{H}$ can be relaxed to the the quantifier "for all h in some orthonormal basis of $\mathcal{H}$".
(iii) The non-degeneracy condition 3 is equivalent to saying that for all smooth $f\neq 0$, the zero-mean Gaussian random variable $\int_0^T f dX$ (which exists as Young integral or via integration-by-parts) has positive definite variance.
(iv) The non-degeneracy condition 3 is equivalent to saying that for all times $0<t_1<...<t_n<T$ the covariance matrix of $(X_{t_1},...,X_{t_n})$, that is,

$$\left(R\left(t_i,t_j\right)\right)_{i,j=1,\dots,d}$$

is (strictly) positive definite.
(v) Non-degeneracy on $[0,T]$ implies non-degeneracy on $[0,t]$ for any $t\in(0,T]$.

Condition 4. ***"0-1 law":*** *The germ σ-algebra $\cap_{t>0}\sigma\left(X_s : s\in[0,t]\right)$ contains only events of probability zero or one.*

When X is Brownian motion, this is the well-known Blumenthal zero-one law. More generally, it holds whenever X is an adapted functional of Brownian motion, including all examples (such as fBM) in which X has a Volterra presentation [8]

$$X_t = \int_0^t K(t,s)\, dB_s.$$

(Nothing is assumed on K other than having the above Wiener-Itô integral well-defined.) The 0-1 law also holds when X is the strong solution of an SDE driven by Brownian motion; this includes the Ornstein-Uhlenbeck - and Browian bridge examples. An example where the 0-1 law fails is given by the *random-ray* $X : t \mapsto tB_T(\omega)$ in which case the germ-event $\{\omega : dX_t(\omega)/dt|_{t=0+} \geq 0\}$ has probability $1/2$. (In fact, sample path differentiability at $0+$ implies non-triviality of the germ σ-algebra see [10] and references therein). We observe that the random ray example is (a) already ruled out by condition 3 and (b) should be ruled out anyway since it does not trigger to the bracket phenomenon needed for a Hörmander statement.

The next condition expresses some sort of scaled support statement at $t = 0+$ and is precisely what is needed in the last part (Step 4) in the proof of the Theorem 1 below. We give examples and easier-to-check conditions below. To state it, we recall [25, Thm 2.2.1] that a geometric p-rough path $\mathbf{x}$ lifts uniquely lifts uniquely and continuously (with respect to homogenous p-variation distances) to a path in the free step-N nilpotent group[5], say

$$S_N(\mathbf{x}) \in C_0^{p\text{-var}}\left([0,T], G^N\left(\mathbb{R}^d\right)\right) \text{ for } N \geq [p].$$

We also recall that $G^N\left(\mathbb{R}^d\right)$ carries a dilation operator δ which generalizes scalar multiplication on $\mathbb{R}^d$.

Condition 5. *Assume there exists $H \in (0,1)$ such that for all fixed $N \geq [p]$, writing $\tilde{\mathbf{X}} = S_N(\mathbf{X})$, all $g \in G^N\left(\mathbb{R}^d\right)$ and for all $\varepsilon > 0$,*

$$\liminf_{n\to\infty} \mathbb{P}\left(d\left(\delta_{n^H}\tilde{\mathbf{X}}_{1/n}, g\right) < \varepsilon\right) > 0.$$

Proposition 3. *Let B denote d-dimensional fractional Brownian motion with fixed Hurst parameter $H \in (1/4, 1)$ and consider the lift to a (random) geometric p-rough path, denoted by $\mathbf{X} = \mathbf{B}$, with $p < 4$. Then it satisfies condition 5.*

Remark 1. *Brownian motion is covered with $H = 1/2$.*

Proof. Write $\tilde{\mathbf{B}} = S_N(\mathbf{B})$. From section 3, and the references therein, the support of the law of $\mathbf{B}$ w.r.t. homogeneous p-variation distance is $C_0^{0,p\text{-var}}\left([0,T], G^{[p]}\left(\mathbb{R}^d\right)\right)$, that is, the closure of lifted smooth path started at 0 with respect to homogeneous p-variation distance [25, 17]. By continuity of S_N [25, Thm 2.2.1] followed by evaluation of the path at time 1 it follows that the support of the law of $\tilde{\mathbf{B}}_1$ is full, that is, equal to $G^N\left(\mathbb{R}^d\right)$. On the other hand, fractional scaling $\left(n^H B_{t/n} : t \geq 0\right) \overset{D}{=} (B_t : t \geq 0)$ implies $\delta_{n^H}\tilde{\mathbf{B}}_{1/n} \overset{D}{=} \tilde{\mathbf{B}}_1$ and so, thanks to full support of $\tilde{\mathbf{B}}_1$,

$$\liminf_{n\to\infty} \mathbb{P}\left(d\left(\delta_{n^H}\tilde{\mathbf{B}}_{1/n}, g\right) < \varepsilon\right) = \mathbb{P}\left(d\left(\tilde{\mathbf{B}}_1, g\right) < \varepsilon\right) > 0.$$

□

[5]The 0 in $C_0^{p\text{-var}}$ indicates that $\mathbf{X}_0$ is started at the unit element in the group.

Although scaling was important in the previous proof, it is only used at times near 0+. One thus suspects that every other Gaussian signal X which scales similarly (on the level of N^{th} iterated integrals!) also satisfies condition 5. To make this precise we need

Theorem 2 ([17]). *Let $(X,Y) = (X^1, Y^1, \ldots, X^d, Y^d)$ be a centered continuous Gaussian process on $[0,1]$ such that (X^i, Y^i) are independent for $i = 1, ..., d$. Let $\rho \in [1,2)$ and assume the covariance of (X,Y), as function on $[0,1]^2$, is of finite ρ-variation (in 2D sense[6]). Then, for every $p > 2\rho$, X and Y can be lifted to geometric p-rough paths denoted $\mathbf{X}$ and $\mathbf{Y}$. Moreover, there exist a constant C depending only on p, ρ, the covariance of (X,Y) so that for all $q \in [1,\infty)$,*

$$\left|d_{p\text{-}var}(\mathbf{X}, \mathbf{Y})\right|_{L^q(\mathbb{P})} \leq C\sqrt{q}\left|R_{X-Y}\right|^{\theta}_{\infty;[0,1]^2}.$$

(Note that $R_{X-Y}(s,t)$ is a diagonal matrix with entries depending on s,t.)

Corollary 1. *Let (X,B) satisfy the conditions of the previous theorem and assume that B is a (d-dimensional) fractional Brownian motion with fixed Hurst parameter $H \in (1/4, 1)$. Assume in addition that*

$$n^{2H}\left|R_{X-B}\right|_{\infty;[0,1/n]^2} \to 0. \tag{4.2}$$

Then condition 5 holds.

Proof. With focus on one diagonal entry and with mild abuse of notation (writing X, B instead of X^i, B^i...)

$$\begin{aligned} & n^{2H}\left|R_{X-B}\right|_{\infty;[0,1/n]^2} \\ = & \sup_{s,t\in[0,1]} \mathbb{E}[n^H\left(X_{s/n} - B_{s/n}\right) n^H \left(X_{t/n} - B_{t/n}\right)] \end{aligned}$$

which can be rewritten in terms of the rescaled process $X^{(n)} = n^H X_{\cdot/n}$, and similarly for B, as

$$\sup_{s,t\in[0,1]} \mathbb{E}\left[\left(X_s^{(n)} - B_s^{(n)}\right)\left(X_t^{(n)} - B_t^{(n)}\right)\right] = \left|R_{X^{(n)}-B^{(n)}}\right|_{\infty;[0,1]^2}.$$

By assumption and the previous theorem, this entails that

$$d_{p\text{-var}}\left(\mathbf{X}^{(n)}, \mathbf{B}^{(n)}\right) \to 0 \text{ in probability.}$$

By continuity of S_N, still writing $\tilde{\mathbf{X}}^{(n)} = S_N\left(\mathbf{X}^{(n)}\right)$ for fixed N, and similarly for $\mathbf{B}^{(n)}$, we have

$$d\left(\tilde{\mathbf{X}}_1^{(n)}, \tilde{\mathbf{B}}_1^{(n)}\right) \leq d_{p\text{-var};[0,1]}\left(\tilde{\mathbf{X}}^{(n)}, \tilde{\mathbf{B}}^{(n)}\right) \to 0 \text{ in probability.}$$

---

[6]Given a function f from $[0,1]^2$ into some normed space, its variation (in the 2D sense!) is an immediate generalization of the standard definition but based on "increments" of form

$$\Delta_{[a,b)\times[c,d)} = f(b,d) + f(a,c) - f(a,d) - f(b,c).$$

But then

$$\begin{aligned}
& \mathbb{P}\left(d\left(\delta_{n^H}\tilde{\mathbf{X}}_{1/n}, g\right) < \varepsilon\right) \\
= & \mathbb{P}\left(d\left(\tilde{\mathbf{X}}_1^{(n)}, g\right) < \varepsilon\right) \\
\geq & \mathbb{P}\left(d\left(\tilde{\mathbf{X}}_1^{(n)}, \tilde{\mathbf{B}}_1^{(n)}\right) + d\left(\tilde{\mathbf{B}}_1^{(n)}, g\right) < \varepsilon\right) \\
\geq & \mathbb{P}\left(d\left(\tilde{\mathbf{B}}_1^{(n)}, g\right) < \varepsilon/2\right) \\
& -\mathbb{P}\left(d\left(\tilde{\mathbf{X}}_1^{(n)}, \tilde{\mathbf{B}}_1^{(n)}\right) > \varepsilon/2\right)
\end{aligned}$$

and so

$$\liminf_{n\to\infty}\mathbb{P}\left(d\left(\delta_{n^H}\tilde{\mathbf{X}}_{1/n}, g\right) < \varepsilon\right) \geq \liminf_{n\to\infty}\mathbb{P}\left(d\left(\tilde{\mathbf{B}}_1^{(n)}, g\right) < \varepsilon/2\right)$$

and this is positive by the example in which we discussed the case of $\mathbf{B}$ resp. $\tilde{\mathbf{B}}$. The proof is finished. □

Example 5 (Ornstein Uhlenbeck)**.** *Given a Brownian motion B, we consider the Ornstein Uhlenbeck (short:OU) process given by the Itô integral*

$$X_t = \int_0^t e^{-(t-r)} dB_r.$$

If B is d-dimensional this yields, component-wise, the d-dimensional OU process. It is readily checked that (X, B) satisifes the assumptions of theorem 2. (In fact, one sees $\rho = 1$ and we are dealing with geometric p-rough paths of Brownian regularity, i.e. $p = 2 + \varepsilon$.) Condition (4.2) then holds with $H = 1/2$: take $s, t \in [0, 1/n]$ and compute, with focus on one non-diagonal entry,

$$\begin{aligned}
R_{X-B}(s,t) & \equiv \mathbb{E}[\,(X_s - B_s)(X_t - B_t)] \\
& = \int_0^t \left(e^{-(s-r)} - 1\right)\left(e^{-(t-r)} - 1\right) dr = O\left(n^{-3}\right).
\end{aligned}$$

By corollary 1 we see condition 5 holds for the Ornstein Uhlenbeck examples.

Example 6 (Brownian Bridge)**.** *Writing $X_t^T := B_t - \frac{t}{T}B_T$ where B is standard Brownian motion this follows along the same lines, again by comparison with B_t for $t \to 0+$.*

5. Taylor Expansions for Rough Differential Equations

Given a smooth vector field W and smooth driving signal $x(\cdot)$ for the ODE $dy = V(y)\,dx$, it follows from (2.2) that

$$J_{0\leftarrow t}^x\left(W\left(y_t^x\right)\right) = W\left(y_0\right) + \int_0^t J_{0\leftarrow s}^x\left(\left[V_i, W\right]\left(y_s^x\right)\right) dx_s^i,$$

where Einstein's summation convention is used throughout. Iterated use of this leads to the Taylor expansion

$$
\begin{aligned}
J^x_{0\leftarrow t}\left(W\left(y^x_t\right)\right) &= W|_{y_0} + \left[V_i, W\right]|_{y_0}\mathbf{x}^{1;i}_{0,t} \\
&\quad + \left[V_i, \left[V_j, W\right]\right]|_{y_0}\mathbf{x}^{2;i,j}_{0,t} \\
&\quad + \ldots \\
&\quad + \left[V_{i_1}, \ldots \left[V_{i_N}, W\right]\right]|_{y_0}\mathbf{x}^{N;i_1,\ldots,i_N}_{0,t} \\
&\quad + \cdots
\end{aligned}
$$

where $\mathbf{x}_{0,t}$ denotes the signature of $x\left(\cdot\right)|^i_{[0,t]}$ in $\mathbb{R}^d \oplus \left(\mathbb{R}^d\right)^{\otimes 2} \oplus \cdots \oplus \left(\mathbb{R}^d\right)^{\otimes N} \oplus \ldots$ Such an expansion makes immediate sense when x is replaced by a geometric p-rough path[7]. Remainder estimates can be obtained via Euler-estimates [19] provided $J^x_{0\leftarrow t}\left(W\left(y^x_t\right)\right)$ is a solution of some ODE of form $dz = \hat{V}\left(z\right)dx$. This is accomplished by setting

$$
z := \left(z^1, z^2, z^3\right) := \left(y^x, J^x_{0\leftarrow t}, J^x_{0\leftarrow t}\left(W\left(y^x_t\right)\right)\right) \in \mathbb{R}^e \oplus \mathbb{R}^{e\times e} \oplus \mathbb{R}^e
$$

Noting that $J^x_{0\leftarrow t}\left(W\left(y^x_t\right)\right)$ is given by $z^2 \cdot W\left(z^1\right)$ in terms of matrix multiplication we have

$$
\begin{aligned}
dz^1 &= V_i\left(z^1\right)dx^i \\
dz^2 &= -z^2 \cdot DV_i\left(z^1\right)dx^i \\
dz^3 &= \left(dz^2\right)\cdot W\left(z^1\right) + z^2 \cdot d\left(W\left(z^1\right)\right) \\
&= z^2 \cdot \left(-DV_i\left(z^1\right)\cdot W\left(z^1\right) + DW\left(z^1\right)\cdot V_i\left(z^1\right)\right)dx^i \\
&= z^2 \cdot \left[V_i, W\right]|_{z^1}dx^i
\end{aligned}
$$

started from $\left(y_0, I, W\left(y_0\right)\right)$ where I denotes the identity matrix in $\mathbb{R}^{e\times e}$ and we see that $\hat{V}$ is given by

$$
\hat{V}_i\left(z^1, z^2, z^3\right) = \begin{pmatrix} V_i\left(z^1\right) \\ -z^2 \cdot DV_i\left(z^1\right) \\ z^2 \cdot \left[V_i, W\right]\left(z_1\right) \end{pmatrix}, \quad i = 1, \ldots, d.
$$

Lemma 2. *Assume $V_1, \ldots, V_d, W$ are smooth vector fields, bounded with all derivatives bounded. Then $\hat{V} = \left(\hat{V}_1, \ldots, \hat{V}_d\right)$ is a collection of smooth (possibly unbounded) vector fields but explosion does not occur. More precisley, there exists a unique RDE solution to $dz = \hat{V}\left(z\right)d\mathbf{x}$ on any compact time interval $[0,T]$. In fact, for some increasing function φ from $\mathbb{R}^+$ into itself*

$$
|z|_{\infty;[0,t]} \leq \varphi\left(M\right) \text{ when } \|\mathbf{x}\|_{p\text{-}var;[0,t]} \leq M.
$$

Proof. Smoothness of $\hat{V}$ is obvious and so the RDE $dz = \hat{V}\left(z\right)d\mathbf{x}$ has a solution up to some possible explosion time. From the particular structure of $\hat{V}$ we now argue that explosion cannot occur in finite time: z^1 does not explode as it is a genuine RDE solution along Lip vector fields, z^2 does not explode as it satisfies a linear RDE (driven by some rough path $M^{\mathbf{x}}$ as already remarked in the proof of Proposition 1). Clearly then, $z^3 = z^2 \cdot W\left(z^1\right)$ where W is a bounded vector fields cannot explode. More precisely, using the estimates for RDE solutions along *Lip* respectively linear vector fields in [19] respectively [25] it is clear that z remains in

[7]By definition, such a p-rough path takes values in the step-$[p]$ tensor algebra but recall that there is a unique lift to the step-N group for any $N > [p]$.

a ball of radius only depending on M if $\|\mathbf{x}\|_{p\text{-var};[0,t]} \leq M$. (With some care one can show that $\log\varphi(M) = O(M^p)$ as $M \to \infty$ but this is irrelevant for the sequel.) $\square$

Let us make the following definitions: given $(m-1)$-times differentiable vector fields $V = (V_1, ..., V_d)$ on $\mathbb{R}^e$, $\mathbf{g} \in \oplus_{k=0}^m (\mathbb{R}^d)^{\otimes k}$ and $y \in \mathbb{R}^e$ we write

$$\mathcal{E}_{(V)}(y, \mathbf{g}) := \sum_{k=1}^m \sum_{\substack{i_1,\dots,i_k \\ \in\{1,\dots,d\}}} \mathbf{g}^{k,i_1,\cdots,i_k} V_{i_1} \cdots V_{i_k} I(y).$$

(Here I denote the identitify function on $\mathbb{R}^e$ and vector fields identified with first order differential operators.) In a similar spirit, given another sufficiently smooth vector field W we first write

$$[V_{i_1}, V_{i_2}, \cdots V_{i_k}, W] := [V_{i_1}, [V_{i_2}, \cdots [V_{i_k}, W] \cdots]]$$

(which may be viewed as first order differential operator) and then

$$\mathbf{g}^k \cdot [V, \dots, V, W]|_{y_0} := \sum_{\substack{i_1,\dots,i_k \\ \in\{1,\dots,d\}}} \mathbf{g}^{k,i_1,\cdots,i_k} [V_{i_1}, V_{i_2}, \cdots V_{i_k}, W] \tag{5.1}$$

with the convention that $\mathbf{g}^0 \cdot V_k = V_k$.

Proposition 4. *(Localized Euler Estimates) Given $\|\mathbf{x}\|_{p\text{-}var;[0,t]} \leq M$ and some integer $m > p-1$ there exists $C = C(M) = C = C\left(M, m, p, \hat{V}\right)$ such that*

$$\left|\pi_{\hat{V}}(z_0, 0; \mathbf{x})_{0,t} - \mathcal{E}_{(\hat{V})}\left(z_0, S_m(\mathbf{x})_{0,t}\right)\right| \leq C(M) \times t^{\frac{m+1}{p}}.$$

Proof. If $\hat{V}$ were bounded with bounded derivatives this would be a consequence of [19, Thm 19]. On the other hand, z must remain in the ball $B(0, \varphi(M))$ and we can replace $\hat{V}$ by (compactly supported) vector fields $\tilde{V}$ such that $\hat{V} \equiv \tilde{V}$ on $B(0, \varphi(M))$. After this localization we apply [19, Thm 19]. $\square$

Lemma 3. *Let f be a smooth function on $\mathbb{R}^e$ lifted to a smooth function on $\mathbb{R}^e \oplus \mathbb{R}^{e\times e} \oplus \mathbb{R}^e$ by*

$$\hat{f}\left(z^1, z^2, z^3\right) = f\left(z^3\right).$$

Viewing vector fields as first order differential operators, we have

$$\hat{V}_{i_1} \cdots \hat{V}_{i_N}|_{z_0} \hat{f} = [V_{i_1}, \cdots, V_{i_N}, W]|_{y_0} f.$$

As a consequence, if I denotes the identity function on $\mathbb{R}^e$,

$$\begin{aligned} & \left| z_t^3 - W|_{y_0} - \sum_{k=1}^m \mathbf{X}_{0,t}^k \cdot [V, \cdots, V, W]|_{y_0} \right| \\ \leq \quad & \left| z_{0,t} - \mathcal{E}_{(\hat{V})}\left(z_0, S_m(\mathbf{x})_{0,t}\right)\right|. \end{aligned}$$

Proof. Taylor expansion of the evolution ODE of $z^3(t)$ shows that $\hat{V}_{i_1} \cdots \hat{V}_{i_N}|_{z_0} f = [V_{i_1}, \cdots, V_{i_N}, W]|_{y_0} f$. $\square$

Corollary 2. *Fix $a \in \mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e$ with $|a| = 1$. Then*

$$\mathbb{P}\left[\left|a^T J_{0\leftarrow t}^x\left(W\left(y_t^x\right)\right) - \sum_{k=0}^m a^T\left(\mathbf{X}_{0,t}^k \cdot [V, \cdots, V, W]|_{y_0}\right)\right|_{t=1/n} > \frac{\varepsilon}{2} n^{-mH}\right] \to 0 \text{ with } n \to \infty.$$

Proof. We estimate this probability by

$$\mathbb{P}\left[\left|a^T J^x_{0\leftarrow t}\left(W\left(y^x_t\right)\right)-\sum_{k=0}^m a^T\left(\mathbf{X}^k_{0,t}\cdot[V,\cdots,V,W]\,|_{y_0}\right)\right|_{t=1/n}>\frac{\varepsilon}{2}n^{-mH};\|\mathbf{x}\|_{p\text{-var};[0,1/n]}\leq 1\right]$$

$$+\mathbb{P}\left[\|\mathbf{x}\|_{p\text{-var};[0,1/n]}\geq 1\right] \qquad \text{then, using } |a|=1 \text{ and the previous lemma,}$$

$$\leq \mathbb{P}\left[\left|z_{0,1/n}-\mathcal{E}_{(\hat{V})}\left(z_0,S_m\left(\mathbf{x}\right)_{0,1/n}\right)\right|>\frac{\varepsilon}{2}n^{-mH};\|\mathbf{x}\|_{p\text{-var};[0,1/n]}\leq 1\right]+o\left(1\right)$$

$$\leq \mathbb{P}\left[C\left(1\right)\times\left(\frac{1}{n}\right)^{\frac{m+1}{p}}>\frac{\varepsilon}{2}n^{-mH}\right]+o\left(1\right) \text{ using the localized Euler estimates.}$$

The probability of the (deterministic) event

$$C\left(1\right)\left(\frac{1}{n}\right)^{\frac{m+1}{p}}>\frac{\varepsilon}{2}\left(\frac{1}{n}\right)^{mH}$$

will be zero for n large enough provided $\frac{m+1}{p}>mH$ which is the case since $p\geq 1,H\leq 1$. □

6. On Hörmander's condition

Let $V=(V_1,...,V_d)$ denote a collection of smooth vector fields defined in a neighbourhood of $y_0\in\mathbb{R}^e$. Given a multi-index $I=(i_1,...,i_k)\in\{1,...,d\}^k$, with length $|I|=k$, the vector field V_I is defined by iterated Lie brackets

$$V_I:=[V_{i_1},V_{i_2},...,V_{i_k}]\equiv[V_{i_1},[V_{i_2},...,[V_{i_{k-1}},V_{i_k}]...]. \tag{6.1}$$

If W is another smooth vector field defined in a neighbourhood of $y_0\in\mathbb{R}^e$ we write[8]

$$\underbrace{a}_{\in(\mathbb{R}^d)^{\otimes(k-1)}}\cdot\underbrace{[V,\ldots,V,W]}_{\text{length }k}:=\sum_{\substack{i_1,\ldots,i_{k-1}\\ \in\{1,\ldots,d\}}}a^{i_1,\ldots,i_{k-1}}[V_{i_1},V_{i_2},...,V_{i_{k-1}},W]$$

Recall that the step-r free nilpotent group with d generators, $G^r(\mathbb{R}^d)$, was realized as submanifold of the tensor algebra

$$T^{(r)}\left(\mathbb{R}^d\right)\equiv\oplus_{k=0}^r\left(\mathbb{R}^d\right)^{\otimes k}.$$

Definition 2. *Given $r\in\mathbb{N}$ we say that condition (H)$_r$ holds at $y_0\in\mathbb{R}^e$ if*

$$\operatorname{span}\left\{V_I|_{y_0}:|I|\leq r\right\}=\mathcal{T}_{y_0}\mathbb{R}^e\cong\mathbb{R}^e; \tag{6.2}$$

Similarly, we say that (HT)$_r$ holds at y_0 if the span of

$$\left\{\pi_{k-1}\left(\mathbf{g}\right)\cdot\underbrace{[V,...,V,V_i]}_{\textit{length } k}|_{y_0}:k=1,...,r;i=1,...,d,\ \mathbf{g}\in G^{r-1}(\mathbb{R}^d)\right\}. \tag{6.3}$$

equals $\mathcal{T}_{y_0}\mathbb{R}^e\cong\mathbb{R}^e$. Hörmander's condtion (H) is satisfied at y_0 iff (H)$_r$ holds for some $r\in\mathbb{N}$. Similarly, we say that the Hörmander-type condtion (HT) is satisfied at y_0 iff (HT)$_r$ holds for some $r\in\mathbb{N}$. (When no confusion arises we omit reference to y_0.)

---

[8]We introduced this notation already in the previous section, cf. (5.1).

Proposition 5. *For any fixed $r \in \mathbb{N}$, the span of (6.2) equals the span of (6.3). Consequently, Hörmander's condition (H) at y_0 is equivalent to the Hörmander-type condtion (HT) at y_0.*

Proof. Given a multi-index $I = (i_1, ..., i_{k-1}, i_k)$ of length $k \leq r$ and writing $e_1, ...e_d$ for the canonical basis of $\mathbb{R}^d$

$$\begin{aligned} \mathbf{g} &= \mathbf{g}(t_1, \ldots, t_{k-1}) \\ &= \exp(t_1 e_{i_1}) \otimes \cdots \otimes \exp\left(t_{k-1} e_{i_{k-1}}\right) \\ &\in G^{r-1}(\mathbb{R}^d) \subset T^{r-1}\left(\mathbb{R}^d\right). \end{aligned}$$

(Recall that $T^{r-1}\left(\mathbb{R}^d\right)$ is a tensor algebra with multiplication $\otimes$, exp is defined by the usual series and the CBH formula shows that the so-defined g is indeed in $G^{r-1}(\mathbb{R}^d)$ as claimed.) It follows that any

$$\pi_{k-1}(\mathbf{g}) \cdot \underbrace{[V, ..., V, V_{i_k}]}_{\text{length } k} |_{y_0}$$

lies in the $(\mathrm{HT})_r$-span i.e. the linear span of (6.3). Now, the $(\mathrm{HT})_r$-span is a closed linear subspace of $\mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e$ and so it is clear that any element of form

$$\pi_{k-1}(\partial_\alpha \mathbf{g}) \cdot \underbrace{[V, ..., V, V_{i_k}]}_{\text{length } k} |_{y_0}$$

where ∂_α stands for any higher order partial derivative with respect to $t_1, ..., t_{k-1}$ i.e.

$$\partial_\alpha = \left(\frac{\partial}{\partial t_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial t_{k-1}}\right)^{\alpha_{k-1}} \quad \text{with } \alpha \in (\mathbb{N} \cup \{0\})^{k-1}$$

is also in the $(\mathrm{HT})_r$-span for any $t_1, \ldots, t_{k-1}$ and, in particular, when evaluated at $t_1 = \cdots = t_{k-1} = 0$. For the particular choice $\alpha = (1, \ldots, 1)$ we have

$$\frac{\partial^{k-1}}{\partial t_1 \ldots \partial t_{k-1}} \mathbf{g}|_{t_1 = 0, \ldots, t_{k-1} = 0} = e_{i_1} \otimes \cdots \otimes e_{i_{k-1}} =: \mathbf{h}$$

where $\mathbf{h}$ is an element of $T^{r-1}\left(\mathbb{R}^d\right)$ with the only non-zero entry arising on the $(k-1)^{th}$ tensor level, i.e.

$$\pi_{k-1}(\mathbf{h}) = e_{i_1} \otimes \cdots \otimes e_{i_{k-1}}.$$

Thus,

$$\pi_{k-1}(\mathbf{h}) \cdot \underbrace{[V, ..., V, V_{i_k}]}_{\text{length } k} |_{y_0} = \left[V_{i_1}, \ldots, V_{i_{k-1}}, V_{i_k}\right] |_{y_0}$$

is in our $(\mathrm{HT})_r$- span. But this says precisely that, for any multi-index I of lenght $k \leq r$ the bracket vector field evaluated at y_0 i.e. $V_I|_{y_0}$ is an element of our $(\mathrm{HT})_r$-span. □

7. Proof of Main Result

We are now in a position to give

Proof (of Theorem 1). We fix $t \in (0, T]$. As usual it suffices to show a.s. invertibility of

$$\sigma_t = \left(\left\langle DY_t^i, DY_t^j \right\rangle_{\mathcal{H}}\right)_{i,j=1,\ldots,e} \in \mathbb{R}^{e \times e}.$$

In terms of an ONB (h_n) of the Cameron Martin space we can write

$$
\begin{aligned}
\sigma_t &= \sum_n \langle DY_t, h_n \rangle_{\mathcal{H}} \otimes \langle DY_t, h_n \rangle_{\mathcal{H}} \\
&= \sum_n \int_0^t J^{\mathbf{X}}_{t \leftarrow s} \left(V_k \left(Y_s \right) \right) dh^k_{n,s} \otimes \int_0^t J^{\mathbf{X}}_{t \leftarrow s} \left(V_l \left(Y_s \right) \right) dh^l_{n,s}
\end{aligned}
\tag{7.1}
$$

(Summation over up-down indices is from here on tacitly assumed.) Invertibility of σ is equivalent to invertibility of the reduced covariance matrix

$$
C_t := \sum_n \int_0^t J^{\mathbf{X}}_{0 \leftarrow s} \left(V_k \left(Y_s \right) \right) dh^k_{n,s} \otimes \int_0^t J^{\mathbf{X}}_{0 \leftarrow s} \left(V_l \left(Y_s \right) \right) dh^l_{n,s}
$$

which has the advantage of being adapted, i.e. being $\sigma \left(X_s : s \in [0, t] \right)$-measurable. We now assume that

$$
\mathbb{P} \left(\det C_t = 0 \right) > 0
$$

and will see that this leads to a contradiction with Hörmander's condition.

Step 1: Let K_s be the random subspace of $\mathcal{T}_{y_0} \mathbb{R}^e \cong \mathbb{R}^e$. spanned by

$$
\left\{ J^{\mathbf{X}}_{0 \leftarrow r} \left(V_k \left(Y_r \right) \right) ; r \in [0, s] , k = 1, ..., d \right\} .
$$

The subspace $K_{0+} = \cap_{s>0} K_s$ is measurable with respect to the germ σ-algebra and by our "0-1 law" assumption, deterministic with probability one. A random time is defined by

$$
\Theta = \inf \left\{ s \in (0, t] : \dim K_s > \dim K_{0+} \right\} \wedge t,
$$

and we note that $\Theta > 0$ a.s. For any vector $v \in \mathbb{R}^e$ we have

$$
v^T C_t v = \sum_n \left| \int_0^t v^T J^{\mathbf{X}}_{0 \leftarrow s} \left(V_k \left(Y_s \right) \right) dh^k_{n,s} \right|^2 .
$$

Assuming $v^T C_t v = 0$ implies

$$
\forall n : \int_0^t v^T J^{\mathbf{X}}_{0 \leftarrow s} \left(V_k \left(Y_s \right) \right) dh^k_{n,s} = 0
$$

and hence, by our non-degeneracy condition on the Gaussian process

$$
v^T J^{\mathbf{X}}_{0 \leftarrow s} \left(V_k \left(Y_s \right) \right) = 0
$$

for any $s \in [0, t]$ and any $k = 1, ..., d$ which implies that v is orthogonal to K_t. Therefore, $K_{0+} \neq \mathbb{R}^e$, otherwise $K_s = \mathbb{R}^e$ for every $s > 0$ so that v must be zero, which implies C_t is invertible a.s. in contradiction with our hypothesis.

Step 2: We saw that K_{0+} is a deterministic and linear subspace of $\mathbb{R}^e$ with strict inclusion $K_{0+} \subsetneqq \mathbb{R}^e$ In particular, there exists a deterministic vector $z \in \mathbb{R}^e \setminus \{0\}$ which is orthogonal to K_{0+}. We will show that z is orthogonal to to all vector fields and (suitable) brackets evaluated at y_0, thereby contradicting the fact that our vector fields satisfy Hörmander's condition. By definition of Θ, $K_{0+} \equiv K_t$ for $0 \leq t < \Theta$ and so for every $k = 1, ...d$,

$$
z^T J^{\mathbf{X}}_{0 \leftarrow t} \left(V_k \left(Y_t \right) \right) = 0 \text{ for } t \leq \Theta. \tag{7.2}
$$

Observe that, by evaluation at $t = 0$, this implies $z \perp \operatorname{span}\{V_1, ..., V_d\} |_{y_0}$.

Step 3: We call an element $\mathbf{g} \in \oplus_{k=0}^{\infty} \left(\mathbb{R}^d \right)^{\otimes k}$ group-like iff for any $N \in \mathbb{N}$,

$$
\left(\pi_0 \left(\mathbf{g} \right), \ldots, \pi_N \left(\mathbf{g} \right) \right) \in G^N \left(\mathbb{R}^d \right) \subset \oplus_{k=0}^{N} \left(\mathbb{R}^d \right)^{\otimes k} .
$$

We now keep k fixed and make **induction hypothesis** $I(m-1)$:

$$\forall \mathbf{g} \text{ group-like}, j \leq m-1: \; z^T \pi_j(\mathbf{g}) \left[V, \cdots, V; V_k\right]\big|_{y_0} = 0.$$

To this end, take the shortest path $\gamma^n : [0, 1/n] \to \mathbb{R}^d$ such that $S_m(\gamma^n)$ equals $\pi_{1,\ldots,m}(\mathbf{g})$, the projection of $\mathbf{g}$ to the free step-m nilpotent group with d generators, denoted $G^m\left(\mathbb{R}^d\right)$. Then

$$|\gamma^n|_{1\text{-var};[0,1/n]} = \|\pi_{1,\ldots,m}(\mathbf{g})\|_{G^m(\mathbb{R}^d)} < \infty$$

and the scaled path

$$h^n(t) = n^{-H} \gamma^n(t), \;\; H \in (0,1)$$

has length (over the interval $[0, 1/n]$) proportional to n^{-H} which tends to 0 as $n \to \infty$. Our plan is to show that

$$\forall \varepsilon > 0 : \liminf_{n\to\infty} \mathbb{P}\left(\left|z^T J^{h^n}_{0\leftarrow 1/n}\left(V_k\left(y^{h^n}_{1/n}\right)\right)\right| < \varepsilon / n^{mH}\right) > 0 \tag{7.3}$$

which, since the event involved is deterministic, really says that

$$\left|n^{mH} z^T J^{h^n}_{0\leftarrow 1/n}\left(V_k\left(y^{h^n}_{1/n}\right)\right)\right| < \varepsilon$$

holds true for all $n \geq n_0(\varepsilon)$ large enough. Then, sending $n \to \infty$, a Taylor expansion and $I(m-1)$ shows that the l.h.s. converges to

$$\left| z^T \underbrace{n^{mH} \pi_m\left(S_m\left(h^n\right)\right)}_{=\pi_m(\mathbf{g})} \cdot \left[V, \cdots, V; V_k\right]\big|_{y_0} \right| < \varepsilon$$

and since $\varepsilon > 0$ is arbitrary we showed $I(m)$ which completes the induction step.

<u>Step 4:</u> The only thing left to show is (7.3), that is, positivity of $\liminf$ of

$$\begin{aligned}
& \mathbb{P}\left(\left|z^T J^{h^n}_{0\leftarrow 1/n}\left(V_k\left(y^{h^n}_{1/n}\right)\right)\right| < \varepsilon / n^{mH}\right) \\
\geq \;& \mathbb{P}\left(\left|z^T J^{\mathbf{X}}_{0\leftarrow \cdot}\left(V_k\left(y_\cdot\right)\right) - z^T J^{h^n}_{0\leftarrow \cdot}\left(V_k\left(y^{h^n}_\cdot\right)\right)\right|_{\cdot = 1/n} < \varepsilon / n^{mH}\right) \\
& - \mathbb{P}\left(\Theta \leq 1/n\right)
\end{aligned}$$

and since $\Theta > 0$ a.s. it is enough to show that

$$\liminf_{n\to\infty} \mathbb{P}\left(\left|z^T J^{\mathbf{X}}_{0\leftarrow \cdot}\left(V_k\left(y_\cdot\right)\right) - z^T J^{h^n}_{0\leftarrow \cdot}\left(V_k\left(y^{h^n}_\cdot\right)\right)\right|_{\cdot = 1/n} < \varepsilon / n^{mH}\right) > 0.$$

Using $I(m-1)$ + stochastic Taylor expansion (more precisely, corollary 2) this is equivalent to show positivity of $\liminf$ of

$$\mathbb{P}\left(\left|z^T \mathbf{X}^m_{0,\cdot}\left[V, \cdots, V; V_k\right] - z^T J^{h^n}_{0\leftarrow \cdot}\left(V_k\left(y^{h^n}_\cdot\right)\right)\right|_{\cdot = 1/n} < \varepsilon / n^{mH}\right).$$

Rewriting things, we need to show positivity of $\liminf$ of

$$\mathbb{P}(|n^{mH} z^T \left[V, \cdots, V; V_k\right] \mathbf{X}^m_{0,1/n} - \underbrace{z^T n^{mH} J^{h^n}_{0\leftarrow 1/n}\left(V_k\left(y^{h^n}_{1/n}\right)\right)}_{\to z^T[V,\cdots,V;V_k]\pi_m(\mathbf{g})}| < \varepsilon)$$

or, equivalently,

$$\mathbb{P}\left(\left|z^T \left[V, \cdots, V; V_k\right]\big|_{y_0}\left(n^{mH} \mathbf{X}^m_{0,1/n} - \pi_m(\mathbf{g})\right)\right| < \varepsilon\right)$$

But this is implied by condition 5 and so the proof is finished. $\square$

Acknowledgment: The second author is partially supported by a Leverhulme Research Fellowship.

References

[1] Baudoin, F.; Hairer, M.: Hoermander's theorem for fractional Brownian motion, to appear in Probab. Theory Relat. Fields (2007).

[2] Bell, D.: The Malliavin calculus. Reprint of the 1987 edition. Dover Publications, Inc., Mineola, NY, 2006

[3] Bismut, J.-M.: Large deviations and the Malliavin calculus. Progress in Mathematics, 45. Birkhäuser Boston, Inc., Boston, MA, 1984

[4] Cass, T.; Friz, P.; Victoir, N.: Non-degeneracy of Wiener functionals arising from rough differential equations (arXiv:0707.0154)

[5] Cattiaux, P.; Mesnager, L.: Hypoelliptic non-homogeneous diffusions, PTRF 123 (2002), 453-483.

[6] Chen, K.T.: Integration of Paths, Geometric Invariants and a Generalized Baker- Hausdorff Formula Ann.of Math. (2), 65:163-178, 1957

[7] Coutin, L., Qian, Z. : Stochastic analysis, rough path analysis and fractional Brownian motion Probab. Theory Related Fields 122 (2002), 108–140

[8] Decreusefond, L.: Stochastic calculus with respect to volterra processes. Ann. Inst. H. Poincaré, 2004

[9] Delbaen, F., Schachermayer, W.: The Mathematics of Arbitrage, Springer (2006).

[10] Dudley, R. M.: Sample Functions of the Gaussian Process, The Annals of Probability, Vol. 1, No. 1. (Feb., 1973), pp. 66-103.

[11] Feyel, D.; de La Pradelle, A.: Curvilinear integrals along enriched paths. Electron. J. Probab. 11 (2006), no. 34, 860–892

[12] Filipovic, D., Teichmann, J.: On the Geometry of the Term structure of Interest Rates, Proceedings of the Royal Society London A 460, 129-167, 2004

[13] Friz, P., Lyons, T., Stroock, D.: Lévy's area under conditioning. Ann. Inst. H. Poincaré Probab. Statist. 42 (2006), no. 1, 89–101

[14] Friz, P., Victoir, N.: Approximations of the Brownian rough path with applications to stochastic analysis. Ann. Inst. H. Poincaré Probab. Statist. 41 (2005), no. 4, 703–724

[15] Friz, P., Victoir, N.: A note on the notion of geometric rough paths. Probab. Theory Related Fields 136 (2006), no. 3, 395–416

[16] Friz, P., Victoir, N.: A Variation Embedding Theorem and Applications, Journal of Functional Analysis, Vol 239 Nr 2 (2006) 631-637

[17] Friz, P., Victoir, N.: Differential Equations Driven by Gaussian Signals I. (arXiv:0707.0313)

[18] Friz, P., Victoir, N.: Differential Equations Driven by Gaussian Signals II. Preprint.

[19] Friz, P., Victoir, N.: Euler Estimates for Rough Differential Equations, to appear in JDE.

[20] Friz, P., Victoir, N.: Multidimensional Stochastic Processes as Rough Paths: Theory and Applications, Cambridge University Press (forthcoming)

[21] Hairer, M., Mattingly, J.C. : Ergodicity of the 2D Navier-Stokes Equations with Degenerate Stochastic Forcing, Annals of Mathematics, vol. 164 (2006) no. 3

[22] Kusuoka, S.: On the regularity of solutions to SDEs.'Asymptotic problem in probability theory: Wiener functionals and asymptotics' ed. K. D. Elworthy and N. Ikeda, Pitman Res. Notes Math. Ser., 284, pp.90–106, Longman Scientific & Technica, 1993

[23] Kusuoka, S.; Stroock, D.: Applications of the Malliavin calculus. III. J. Fac. Sci. Univ. Tokyo Sect. IA Math. 34 (1987), no. 2, 391–442

[24] Hairer, M.: Ergodicity of stochastic differential equations driven by fractional Brownian motion, Ann. Probab. 33 (2005), no 3, pp. 703-758

[25] Lyons, T.: Differential equations driven by rough signals, Rev. Mat. Iberoamericana 14 (1998), no. 2, 215–310.

[26] Lyons, T.; Qian, Z.: Calculus of variation for multiplicative functionals, New trends in stochastic analysis (Charingworth, 1994), 348-374, World Sci. Publishing, River Edge, NJ, 1997

[27] Lyons, T.; Qian, Z.: Qian, Zhongmin Flow of diffeomorphisms induced by a geometric multiplicative functional. Probab. Theory Related Fields 112 (1998), no. 1, 91–119

[28] Ledoux, M.; Qian, Z.; Zhang, T. Large deviations and support theorem for diffusion processes via rough paths. Stochastic Process. Appl. 102, no. 2, 265–283. 2002

[29] Lyons, T.; Qian, Z.: System Control and Rough Paths, Oxford University Press, 2002.

[30] Malliavin, P.: Stochastic analysis. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 313. Springer-Verlag, Berlin, 1997

[31] Mandelbrot, B. B.; Van Ness, John W. Fractional Brownian motions, fractional noises and applications. SIAM Rev. 10 1968 422–437

[32] Nualart, D.: The Malliavin calculus and related topics. Second edition. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006

[33] Nualart, D.; Saussereau, B.: Malliavin calculus for stochastic differential equations driven by a fractional Brownian motion. Preprint (2005).

[34] Hu, Y.; Nualart, D.: Differential equations driven by Hoelder continuous functions of order greater than 1/2; ArXiv (math.PR/0601628)

[35] Picard, J.: A tree approach to p-variation and to integration. (arXiv:0705.2128)

[36] Shigekawa, I.: Stochastic analysis. Translated from the 1998 Japanese original by the author. Translations of Mathematical Monographs, 224. Iwanami Series in Modern Mathematics. American Mathematical Society, Providence, RI, 2004.

[37] Stroock, D. W.: The Malliavin calculus and its application to second order parabolic differential equations. I. Math. Systems Theory 14 (1981), no. 1, 25–65

[38] Taniguchi, S.: Applications of Malliavin's calculus to time-dependent systems of heat equations. Osaka J. Math., 22 (1985), 307-320.